AD-769 169

A NEW TECHNIQUE IN THE OPTIMIZATION OF EXPONENTIAL QUEUEING SYSTEMS

Steven A. Lippman

California University

Prepared for:

Air Force Office of Scientific Research

October 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AD-769 169

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Western Management Science Institute
University of California, Los Angeles

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

A New Technique in the Optimization of Exponential Queueing Systems

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

Steven A. Lippman

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1973	54	24

8a. CONTRACT OR GRANT NO. AFOSR-72-2349a

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Working Paper No. 211

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Available on [redacted]

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

We consider the problem of controlling M/M/c queueing systems with $c \geq 1$. By providing a new definition of the time of transition, we enlarge the standard set of decision epochs, and obtain a preferred version of the n-period problem in which the times between transitions are exponential random variables with constant parameter. Using this new technique, we are able to utilize the inductive approach in a manner characteristic of inventory theory. The efficacy of the approach is then demonstrated by successfully finding the form of an optimal policy for four quite distinct models that have appeared in the literature; namely, those of (i) McGill, (ii) Miller-Cramer, (iii) Crabill-Sabeti, and (iv) Low. Of particular note, our analysis establishes thatan (s,S) or control-limit policy is, as previously conjectured, optimal for an M/M/c queue with switching costs and removable servers. In addition, it is shown for the Miller-Cramer model that a policy optimal for all sufficiently small discount factors can be obtained from the usual average cost functional equation without recourse to further computation.

DD FORM 1473 1 NOV 65

Security Classification

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
semi-markov decision processes						
dynamic and control of queueing systems						
planning horizons						

Security Classification

Working Paper #211

Operations Research Study Center
Graduate School of Management
University of California, Los Angeles

A NEW TECHNIQUE IN THE OPTIMIZATION OF EXPONENTIAL QUEUEING SYSTEMS *

Steven A. Lippman

October, 1973

*This research was supported by the Air Force Office of Scientific Research through Grant AFOSR-72-2349a.

ABSTRACT

We consider the problem of controlling M/M/c queueing systems with $c \geq 1$, By providing a new definition of the time of transition, we enlarge the standard set of decision epochs, and obtain a preferred version of the n-period problem in which the times between transitions are exponential random variables with constant parameter. Using this new technique, we are able to utilize the inductive approach in a manner characteristic of inventory theory. The efficacy of the approach is then demonstrated by successfully finding the form of an optimal policy for four quite distinct models that have appeared in the literature; namely, those of (i) McGill, (ii) Miller-Cramer, (iii) Crabill-Sabeti, and (iv) Low. Of particular note, our analysis establishes that an (s,S) or control-limit policy is, as previously conjectured, optimal for an M/M/c queue with switching costs and removable servers. In addition, it is shown for the Miller-Cramer model that a policy optimal for all sufficiently small discount factors can be obtained from the usual average cost functional equation without recourse to further computation.

I. INTRODUCTION

Recently, Prabhu and Stidham [19] presented an excellent synthesis and survey of the literature on the optimal control of queueing systems. There, the authors clearly articulated the need for effecting a unified treatment, if not a unified theory, of the optimal control of queueing systems in contrast to the ad hoc manner that has characterized development within the field to date.

With a view towards the goal of providing a unified treatment, a new definition of a transition for exponential or Markovian queueing systems (that is, systems with Poisson arrivals and exponential service times) is introduced in order to facilitate the use of the inductive approach on the finite horizon problems in attempting to specify the form of an optimal policy. Quite simply, this new definition merely stipulates that the exponential holding times between transitions (which normally entail a change in state) have constant parameter. Thus, the times between transitions are independent of not only the control policy employed but also the state of the system. As will be demonstrated, this extraordinarily simple device yields both aesthetic and pragmatic benefits.

As to the aesthetic benefits, we claim that the version of the n-period problem as defined herein is intrinsically more meaningful than the n-period problem induced by the standard definition of a transition for it corresponds to a finite horizon problem of some given expected duration.

Pragmatically, the new definition enables us to readily obtain many new results while simultaneously extracting as a byproduct a number of those that have, with considerable ingenuity and difficulty, been previously established. This is done by applying the inductive approach to show that various functions of the n-period return function are convex and/or monotone.

In particular, four distinct queueing models are considered. First, and foremost, is McGill's [16] M/M/c system with removable servers. Here we established -- for the first time -- the optimality of an (s,S) or control-limit policy for the finite and infinite horizon problems, both with and without discounting. It is shown that if $n \leq +\infty$ periods remain, there are i customers in the system, and the number of servers "on" is not between the boundaries $s_{n,i}$ and $S_{n,i}$, then turn on or off just enough servers to reach the boundary; otherwise, do not change the number of servers on.

The various researchers who have investigated the other three models we consider all adopted maximization of the long-run average expected reward per unit time as their criterion of optimality. As in the example of the M/M/c queue with removable servers, we establish the form of an optimal policy for the finite and infinite horizon problems, both with and without discounting. Moreover, in all three of the models we establish the existence of very strong planning horizons. These three models referred to include: (i) Cramer's [7] extension of Miller's [18] M/M/c system with finite queue capacity in which the customers are distinguished by the reward associated with their acceptance into the queue; (ii) Crabill [5,6] and Sabeti's [20] M/M/1 system in which the server can operate at any one of a finite number of service rates; there is a higher cost associated with faster service rates and a linear holding cost in addition to a reward for service completions, (iii) Low's [13,14] M/M/c finite capacity queue in which the decision maker (in effect) dynamically selects the customer arrival rate so as to balance the linear holding costs against the higher entrance fees that accompany slower arrival rates. This approach will not prove beneficial in all instances; in particular, the approach has not proved useful in analyzing Cramer's models M_3 and M_4 [7 , pp. 38-57].

In summary, the purpose of this paper is twofold. First, a redefinition of transitions is proposed for exponential queueing systems in order to bring into being both a more meaningful and a technically more useful version of the n-period problem and, simultaneously, in order to achieve some small unification in the treatment of exponential queueing systems. The second purpose is to establish new results for four specific models which occupy a rather prominent position in the field vis-a-vis other models to be found in the literature.

The four models are considered in sections 3 through 6 while notation and a more detailed explanation of our new definition of a transition are presented in section 2.

II. THE NEW SET OF DECISION EPOCHS

In pursuing our investigation of the optimal control of exponential queueing systems, presentation of our approach is facilitated by temporarily considering a larger class of problems known as semi-Markov decision processes (SMDP). We begin by defining a SMDP and use the M/M/c queue with removable server to provide a concrete illustration of our many definitions.

A SMDP is specified by five objects: a state space S, an action space $A = X_{s\varepsilon S} A_s$, a law of motion q, a transition time t, and a reward r. Whenever (and however) the system is in state s and we choose action a, three things happen: (1) the system moves to a new state selected according to the probability distribution $q(\cdot|s,a)$, (2) conditional on the event that the new state is s', the length of time it takes the system to move to state s' is a nonnegative random variable with probability distribution $t(\cdot|s,a,s')$, and (3) conditional on the event that the new state is s' and the transition takes t time units, we receive a reward $r(\tau,t|s,a,s')$ by time $\tau \leq t$; typically, the reward is composed of a continuous (as a function of τ) component and a jump which is received either at $\tau = 0$ or at $\tau = t$. After the transition to s' occurs, a new action $a'\varepsilon A_{s'}$ is chosen and the process continues in the obvious manner. The decision epochs are the times of transition.

In the context of a SMDP, we shall assume that the time between transitions is an exponential random variable with parameter $\lambda_{s,a}$; that is, $t(\tau|s,a,s') = 1-e^{-\lambda_{s,a}\tau}$. Moreover, we require that there be a finite upper bound on the set of λ's to ensure that only a finite number of transitions will occur in a finite amount of time.

Letting $\alpha \geq 0$ be the interest rate used for discounting, so that a reward r received at time t has present value $r\,e^{-\alpha t}$, we define $V_{n,\alpha}(s)$ to be the total expected α-discounted reward that can be obtained during the last n transitions when starting from state s and following an optimal policy. (When it is clear that the value of α is fixed, we will often delete the α and simply write $V_n(s)$ rather than $V_{n,\alpha}(s)$.) Setting $V_{0,\alpha}(s) \equiv 0$, it is clear that we have the following recursive equations for $V_{n,\alpha}$:

$$(1) \qquad V_{n+1,\alpha}(s) = \max_{a\varepsilon A_s} \{r_\alpha(s,a) + \int_S \lambda_{s,a}(\alpha+\lambda_{s,a})^{-1} V_{n,\alpha}(s')dq(s'|s,a)\} ,$$

where

$$r_\alpha(s,a) \equiv \int_S\{\int_0^\infty[\int_0^t e^{-\alpha\tau}dr(\tau,t|s,a,s')]\lambda_{s,a}e^{-\lambda_{s,a}t}dt\}dq(s'|s,a) .$$

Of course, $r_\alpha(s,a)$ is the expected α-discounted reward earned during one transition when starting from state s and choosing action a.

If $\langle V_{n,\alpha}(s)\rangle$ possesses a limit for each s as n tends to infinity, then we denote this limit by $V_\alpha(s)$. Throughout the remainder of the paper, we will refer to any policy π whose α-discounted return U_α equals V_α as <u>α-optimal</u>, $\alpha>0$. If $U_\alpha - V_\alpha$ goes to zero as α goes to zero, then π is said to be <u>0-optimal</u>; furthermore, if for some $\alpha'>0$ we have $U_\alpha = V_\alpha$ for $0<\alpha\leq\alpha'$, then π is termed <u>strongly optimal</u>. Finally, π is said to be <u>average optimal</u> if $\liminf_{n\to\infty} U_{n,0}/n = \limsup_{n\to\infty} V_{n,0}/n \equiv \bar{V}$.

As a concrete illustration of the above, consider the M/M/c queueing system with removable servers [16]. Customers arrive according to a Poisson process with rate λ. There are $c < \infty$ independent exponential servers each with rate μ, and the queue has an unlimited capacity. The cost structure consists of three parts: a holding cost h per customer per unit time, a running cost r per server per unit time, and a switching cost $K^+(K^-)$ that is incurred each time a server is turned on (off). Thus, if there are x servers on and it is decided to have y servers on, then the switching cost is given by

$$K(x,y) = \begin{cases} K^+(y-x), & \text{if } y \geq x \\ K^-(x-y), & \text{if } y \geq x . \end{cases}$$

Taking the state of the system to be (i,x), where i is the number of customers in the system and x is the number of servers on, we have

$$S = \{(i,x) : i = 0,1,2,\ldots; x = 0,1,2,\ldots,c\}, \quad A_s \equiv A = \{0,1,2,\ldots,c\} ,$$

and

$$q((i+1,y)\,|\,(i,x),y) = \frac{\lambda}{\lambda+(i\wedge y)\mu} \quad \text{and} \quad q((i-1,y)\,|\,(i,x),y) = \frac{(i\wedge y)\mu}{\lambda+(i\wedge y)\mu} .$$

Of course, $\lambda_{(i,x),y} = \lambda + \mu(i\wedge y)$ so that the holding times depend upon the state (through i) and the action chosen. Finally, $r(\tau,t\,|\,(i,x),y,(i',y)) = K(x,y) + (hi+ry)\tau$, so that $r_\alpha((i,x),y) = K(x,y) + (hi+ry)/(\alpha+\lambda+(i\wedge y)\mu)$.

The most obvious disadvantage of the problem formulation as given above is that the expected length of the n-period problem is not a constant (for fixed n), but rather a variable that depends heavily upon both the control policy implemented and the initial state of the system.

(Roughly speaking, more customers in the system and more servers on will result in a shorter time till the completion of the n^{th} transition.) Thus, the n-period problem does not correspond in any strong sense to a problem in which one seeks to minimize the expected discounted costs over some time horizon of finite (expected) length T. More sharply, we see that the n-period problem with the standard formulation is not, as one would naturally presume it to be, a continuous time analog of the discrete time n-period problem where the periods are of constant and equal length.

While the period lengths in a SMDP must necessarily be random variables, the problem can be reformulated so that the exponential period lengths all have the same parameter, independent of both the control policy employed and the initial state. Hence, the aesthetic-philosophical need to more closely approximate a problem of fixed length -- or at least of fixed expected length -- leads us to advocate the necessity of a reformulation.

A second reason prompting us to reformulate the n-period problem is the fact that the standard formulation dissipates desirable properties -- such as monotonicity and concavity -- of the return function $V_{n,\alpha}(\cdot)$. In turn, this leads to "foolish decisions being optimal." For instance, if $K^{-} > r/\lambda$ and $\mu/\lambda > h/r$ in the removable server model, then $V_{1,\alpha}((1,1)) < V_{1,\alpha}((0,1))$ for all $\alpha \geq 0$ small, so it is cheaper to incur the holding cost simply in order to increase the transition rate and hasten the end of the horizon. On the other hand, it is shown in section 3 that $V_{\alpha}(\cdot)$ is strictly increasing in the number of customers in the system, so that the "improper" standard formulation has lost the monotonicity of $V_{n,\alpha}((\cdot,x))$. Similar aberrations can occur in the policy itself.

We now show how to change the set of decision epochs so as to obtain constant λ's while leaving the underlying stochastic process unchanged. To begin, assume for simplicity in presentation that $q(s|s,a) \equiv 0$, and define

$$\Lambda \equiv \sup_{s,a} \lambda_{s,a} .$$

Next, redefine the law of motion q as follows:

$$q'(s'|s,a) = \begin{cases} \dfrac{\Lambda-\lambda_{s,a}}{\Lambda}, & \text{if } s' = s \\ \dfrac{\lambda_{s,a}}{\Lambda} q(s'|s,a), & \text{if } s' \neq s . \end{cases}$$

This change in the law of motion permits us to redefine the λ's by setting them all equal to Λ. Taken together these changes in q and the λ's leave the underlying stochastic process unchanged as the infinitesimal generator is unchanged and it uniquely determines the pure jump Markov process [4 , Ch. 8].

Having increased $\lambda_{s,a}$ to Λ, the expected length of the time until a transition occurs has been reduced from $1/\lambda_{s,a}$ to $1/\Lambda$. Of course, the number of transitions until a change of state occurs has changed from the constant 1 to a geometric random variable with parameter $\lambda_{s,a}/\Lambda$. Combining these two facts shows (again) that the expected time till a change of state occurs is $[\lambda_{s,a}/\Lambda]^{-1}/\Lambda = 1/\lambda_{s,a}$ as desired.

In the context of the M/M/c queue with removable servers, we have

$$\Lambda = \lambda + c\mu \quad \text{and} \quad r_\alpha((i,x),y) = K(x,y) + (hi+ry)/(\alpha+\Lambda) ,$$

so that the recursive equations for the reformulated problem are

$$(2) \quad V_{n+1}(i,x) = \min_y \{K(x,y) + \frac{1}{\alpha+\Lambda} [hi + ry + \lambda V_n(i+1,y) + \mu(i\wedge y)V_n(i-1,y) + (\Lambda-\lambda-\mu(i\wedge y))V_n(i,y)]\} .$$

As originally conceived, reformulation of this model entailed allowing idle servers to complete service on fictitious customers at the rate μ with the proviso that while we say that a transition has occurred if an idle server completes service, the state of the system does not change. More generally, one can imagine a bell that is triggered by an exponential clock with parameter Λ . A transition occurs if and only if the bell rings. Furthermore, the probability that the new state is s' given that the system was in state s and action a was chosen is given by $q'(s'|s,a)$ and is determined independently of the clock and of the past choices of s'.

In closing, we note that if a stationary policy is employed for the infinite horizon problem, then both the new and the standard formulation are equivalent as are their functional equations. Moreover, for every policy -- Markovian or not -- in the standard formulation, there corresponds a policy in the new formulation with the same sample paths (although not the same system history) and the same return function. Consequently, establishing the existence of a stationary policy that is optimal in the new formulation yields the same result for the standard formulation.

The efficacy of our new formulation is demonstrated in the next four sections.

III. THE M/M/c QUEUE WITH REMOVABLE SERVERS

In this section we consider the M/M/c queue with removable servers described in section 2. Using the inductive approach on the n-period problem with constant expected time between transitions, we begin by showing that the n period return function $V_{n,\alpha}(i,x)$ as given in (2) is a convex function of x. From this, the optimality of a control-limit policy follows readily.

Define $J_{n,\alpha}(i,y)$ by

(3) $$J_{n,\alpha}(i,y) = \frac{1}{\alpha+\Lambda}\{h\,i + r\,y + \lambda V_{n,\alpha}(i+1,y) + \mu(i\wedge y)V_{n,\alpha}(i-1,y) + \mu(c-(i\wedge y))V_{n,\alpha}(i,y)\}$$

so that $(V_{0,\alpha}\equiv 0)$

(4) $$V_{n+1,\alpha}(i,x) = \min_{y}\{K(x,y) + J_{n,\alpha}(i,y)\}.$$

Also, define $y_{n,\alpha}(i,x)$ to be the optimal decision (number of servers on) when n periods remain, the discount factor is α, and the current state is (i,x). We say that $y_{n,\alpha}$ is a control-limit policy if there are integers $s_{n,\alpha}(i)$ and $S_{n,\alpha}(i)$ with $s_{n,\alpha}(i) \le S_{n,\alpha}(i)$ such that

$$y_{n,\alpha}(i,x) = \begin{cases} s_{n,\alpha}(i), & x \le s_{n,\alpha}(i) \\ x\,, & s_{n,\alpha}(i)<x<S_{n,\alpha}(i) \\ S_{n,\alpha}(i), & x \ge S_{n,\alpha}(i). \end{cases}$$

If, in addition, $s_{n,\alpha}(i) \le i$ for each i, then we say that $y_{n,\alpha}$ is a regular control-limit policy.

THEOREM 1: Given $\alpha \geq 0$, n, and i, the functions $V_{n,\alpha}(i,\cdot)$ and $J_{n,\alpha}(i,\cdot)$ are convex and $y_{n,\alpha}$ is a regular control-limit policy.

Proof: Defining $H_n(x,y) = K(x,y) + J_{n,\alpha}(i,y)$, we note that $H_n(\cdot,\cdot)$ is jointly convex as $J_{0,\alpha}(i,\cdot)$ is linear, $K(\cdot,\cdot)$ is jointly convex, and the sum of convex functions is convex. Hence, $V_{1,\alpha}(i,x) = \min_y \{H_0(x,y)\}$ is a convex function of x. (Note that $H_n(\cdot,\cdot)$ convex in each variable is not sufficient; joint convexity is needed to apply this well known theorem on the minimum of a convex function.) Now assume that $V_{n,\alpha}(i,\cdot)$ is convex for each i. Then $J_{n,\alpha}(i,\cdot)$ is convex for each i as the sum of convex functions is convex. Consequently, $H_n(\cdot,\cdot)$ is jointly convex so that $V_{n+1,\alpha}(i,\cdot)$ is convex for each i.

The convexity of $J_{n,\alpha}(i,\cdot)$ coupled with the essentially linear form of $K(\cdot,\cdot)$ yields the existence of $s_{n,\alpha}(i) \leq S_{n,\alpha}(i)$. To see that $s_{n+1,\alpha}(i) \leq i$, simply note that $J_{n,\alpha}(i,i) \leq J_{n,\alpha}(i,i+j) + K^+j/(\alpha+\Lambda)$, so that if action i+j were optimal from state (i,i) we would have

$$V_{n+1,\alpha}(i,i) = K^+j + J_{n,\alpha}(i,i+j) > J_{n,\alpha}(i,i) \geq V_{n+1,\alpha}(i,i).$$

Q.E.D.

Theorem 1 can now be applied to yield the same results for the infinite horizon problem with $\alpha>0$.

THEOREM 2: For $\alpha>0$, $V_\alpha(i,\cdot)$ and $J_\alpha(i,\cdot)$ exist and are convex, $V_\alpha(\cdot,\cdot)$ is the unique solution to the functional equation of dynamic programming, and there is a stationary policy y_α that is α-optimal. Moreover, $y_\alpha \equiv \langle s_\alpha(i), S_\alpha(i)\rangle$ is a regular control-limit policy.

Proof: To see that $V_\alpha \equiv \lim_{n\to\infty} V_{n,\alpha}$ exists, simply note that $V_{n+1} \geq V_n$ and that

$$V_{n,\alpha}(i,x) \leq \sum_{j=0}^{n-1} [K^+c + K^-c + \frac{rc + (i+j)h}{\alpha+\Lambda}]\left(\frac{\Lambda}{\alpha+\Lambda}\right)^j < B_i < \infty.$$

(Of course, $\{B_i\}$ is not bounded.) The convexity of $V_\alpha(i,\cdot)$ follows from Theorem 1 and the fact that the limit of convex functions is convex, whereas the uniqueness and the existence of an optimal stationary policy is immediate from Theorem 1 of [11]. These last facts, coupled with the convexity of $J_n(i,\cdot)$, suffice to establish the existence of $\langle s_\alpha(i)\rangle$ and $\langle S_\alpha(i)\rangle$.

Q.E.D.

Establishing these results for the average cost case is slightly more delicate and, in particular, we will need to assume $\lambda<c\mu$ least $\overline{V}$, the optimal return function, be infinite. Before proving that a control-limit policy is also optimal for the average cost problem, we need the following lemma which asserts that all servers must be on when many periods remain, many customers are in the system, and $\alpha>0$ is small.

LEMMA 1: There are numbers $N^*<\infty$, $I<\infty$, and $\alpha^*>0$ such that $s_{n,\alpha}(i) = c$ whenever $n \geq N^*$, $i \geq I$, and $\alpha \leq \alpha^*$.

Proof: To begin, define $v_{n,\alpha}(i) = \min_x \{V_{n,\alpha}(i+1,x) - V_{n,\alpha}(i,x)\}$. Letting $\xi = y_{n+1,\alpha}(i+1,x)$, we have

$$V_{n+1,\alpha}(i+1,x) - V_{n+1,\alpha}(i,x) \geq \frac{1}{\alpha+\Lambda}\{h + J_{n,\alpha}(i+1,\xi) - J_{n,\alpha}(i,\xi)\}$$

$$\geq \frac{1}{\alpha+\Lambda}\{h + \lambda v_{n,\alpha}(i+1) + \mu(\xi\wedge i)v_{n,\alpha}(i-1) + \mu(c-(\xi\wedge i))v_{n,\alpha}(i)\},$$

so that

$$v_{n+1,\alpha}(i) \geq \frac{h}{\alpha+\Lambda} + \frac{\Lambda}{\alpha+\Lambda}\min\{v_{n,\alpha}(i-1), v_{n,\alpha}(i), v_{n,\alpha}(i+1)\}.$$

Iterating this inequality gives

(5) $$v_{n+1,\alpha}(i+1) \geq \frac{h}{\alpha+\Lambda}\sum_{j=0}^{n\wedge i}\left(\frac{\Lambda}{\alpha+\Lambda}\right)^j$$

Let $\mathcal{V}_{n,\alpha}$ be the return associated with choosing action c when n periods remain and acting in an optimal fashion thereafter. Then for $i \geq c$,

$$V_{n+1,\alpha}(i+1,x) - \mathcal{V}_{n+1,\alpha}(i+1,x) \geq -[cK^+ + \frac{cr}{\alpha+\Lambda}] + \frac{1}{\alpha+\Lambda}\{\lambda[V_{n,\alpha}(i+2,\xi) - V_{n,\alpha}(i+2,c)]$$

$$+ \mu\xi[V_{n,\alpha}(i,\xi) - V_{n,\alpha}(i,c)] + \mu(c-\xi)[V_{n,\alpha}(i+1,\xi) - V_{n,\alpha}(i,c)]\}$$

$$\geq -cK^+ - \frac{cr}{\alpha+\Lambda} - \frac{\Lambda}{\alpha+\Lambda}(c-\xi)K^- + \mu(c-\xi)v_{n,\alpha}(i)$$

$$\geq -c[K^+ + K^- + \frac{r}{\Lambda}] + \mu(c-\xi)v_{n,\alpha}(i).$$

But (5) permits us to choose N^*, α^*, and I so that the right hand side of the inequality above is strictly positive for all $n \geq N^*$, $i \geq I$, and $\alpha \leq \alpha^*$ whenever $\xi < c$. Hence, we must have $y_{n+1,\alpha}(i,x) = \xi = c$ for all $n \geq N^*$, $i \geq I$ and $\alpha \leq \alpha^*$.

Q.E.D.

Note that Equation 5 allows explicit computation of the numbers N^*, I, and α^*.

Because the optimal return function $\overline{V}$ turns out to be constant, another function to play the role of V_α is needed. Toward this end, define the functions h and J by

$$h(i,x) = \lim_{\alpha \to 0^+} \{V_\alpha(i,x) - V_\alpha(0,0)\}$$

and

$$J(i,y) = \frac{1}{\Lambda}\{hi + ry + \lambda h(i+1,y) + \mu(i \wedge y)h(i-1,y) + \mu(c-(i \wedge y))h(i,y)\}.$$

<u>THEOREM 3</u>: If $\lambda < c\mu$, then $\overline{V}$ is constant and finite, h exists and is finite, $h(i,\cdot)$ and $J(i,\cdot)$ are convex for each i, and there is a stationary regular control-limit policy $\overline{y} \equiv \langle(s(i),S(i)\rangle$ with $s(i) = c$ for all i sufficiently large that is average optimal. Furthermore, h satisfies the functional equation

$$h(i,x) = \min_y \{K(x,y) + J(i,y)\} - \overline{V}/\Lambda , \tag{6}$$

and any stationary policy that selects an action which minimizes the right side of (6) for each $s \in S$ is average optimal.

<u>Proof</u>: In view of Lemma 1, the set of policies that is α-optimal for $\alpha \leq \alpha^*$ is finite. Consequently, there is a sequence $\langle\alpha_m\rangle$ of discount factors with $\alpha_m \to 0$ and some control-limit policy $\overline{y}$ that is α_m-optimal for each m. Furthermore, the lower control-limits $\langle s(i)\rangle$ satisfy

$s(i) = c$ for all $i \geq I$. Thus, $\bar{y}$ satisfies the hypotheses of Corollary 1 of [11] so that $\bar{y}$ is average optimal and $\bar{V}$ is constant and finite.

The convexity of h and J follows from the fact that the limit of convex functions is convex while the remainder of the results follow from Theorem 4 of [11].

Q.E.D.

In a forthcoming paper [9], it is shown that there is a strongly optimal control-limit policy. The paper will also contain a further characterization of the parameters $s_{n,\alpha}(i)$ and $S_{n,\alpha}(i)$.

IV. OPTIMAL CUSTOMER SELECTION IN AN M/M/c QUEUE

One of the earliest papers to appear in the literature of the optimal dynamic control of queueing systems was Miller's [17]. He considered an M/M/c system with m customer classes in which each server had rate μ, customer arrivals had rate λ, and p_k was the probability that an arriving customer was from the k^{th} class. There was a reward r_k, $r_1<r_2<\cdots<r_m$ associated with serving a customer of class k. Decisions were made at the times of arrival whence the customer was either accepted into service in order to obtain the reward r_k or rejected in order to keep available servers free. Pre-emption and backlogging (queueing) of customers was not allowed, and maximization of the expected reward earned per unit time over an infinite planning horizon was the criterion of optimality. Several years later, Cramer [7] improved upon this model by introducing a finite queue capacity and allowing an infinite number of customer classes.

The treatment of the model presented here represents a very slight generalization of Cramer's model in that we allow the queue capacity Q to be either finite or infinite. In addition, we consider discounting and finite horizon problems.

To begin, let the set $\mathfrak{X}$ of customer classes be a measurable subset of the interval $[1,K]$, and assume that the reward function $r:\mathfrak{X}\to\mathbb{R}$ is a strictly increasing function with $r_1>0$ and $r_K<\infty$. Denote by p the measure on $\mathfrak{X}$; that is, if $\mathfrak{X}$ were countable, then p_x is the probability that the next arrival will be a customer of class x. Finally add an artificial class 0 to $\mathfrak{X}$ with $r_0=0$ and $p_0=0$.

Let the state of the system be (i,x), where x is the class of the customer seeking admittance and i is the number of customers in the system not including the one seeking admittance, and denote acceptance by 1 and rejection by 0. Then in accord with our previous notation we have

$$S = \{(i,x): i=0,1,2,\cdots,c+Q, x\varepsilon\{0\}\cup\underline{X}\}$$

$A_{(i,x)} = \{0\}$ for $i=c+Q$ or $x=0$ and $A_{(i,x)} = \{0,1\}$ otherwise,

$$q((i,y)|(i,x),0) = q((i+1,y)|(i,x),1) = \frac{\lambda p(dy)}{\Lambda}$$

$$q((i-1,0)|(i,x),0) = \frac{\mu(i\wedge c)}{\Lambda}, \quad q((i,0)|(i,x),1) = \frac{\mu((i+1)\wedge c)}{\Lambda},$$

$$q((i,0)|(i,x),0) = \frac{\mu(c-(i\wedge c))}{\Lambda}, \text{ and } q((i+1,0)|(i,x),1) = \frac{\mu[c-((i+1)\wedge c)]}{\Lambda}.$$

We also ave $\Lambda = \lambda + c\mu$, $r_\alpha((i,x),1) = r_x$ and $r_\alpha((i,x),0) = 0$, so that the recursive equations for this model are $(V_0 \equiv 0)$

(7) $$V_{n+1,\alpha}(i,x) = \max\{r_x + V_{n,\alpha}(i+1); V_{n,\alpha}(i)\}, \quad i<c+Q,\ x\neq 0,$$

where

(8) $$V_{n,\alpha}(i) = \frac{1}{\alpha+\Lambda}\{\lambda\int_{\underline{X}} V_{n,\alpha}(i,y)p(dy) + (i\wedge c)\mu V_{n,\alpha}(i-1,0) + (c-(i\wedge c))\mu V_{n,\alpha}(i,0)\}.$$

Our formulations implicitly assumes that the fee r_x is collected at the time the customer enters the queue rather than at the time he commences service. To handle this latter formulation, simply set

$$r_\alpha((i,x),1) = \left[\frac{c\mu}{\alpha+c\mu}\right]^{(i+1-c)\vee 0} r_x \tag{9}$$

as the customer will not enter service until $(i+1-c)$ other customers complete service and each of these independent service times has parameter c . Of course, when $Q=\infty$, the problem is of interest only if r_α is given by (9). Because the case $Q=\infty$ is so different from the case $Q<\infty$ (e.g., $V_{n,\alpha}(\cdot,x)$ is concave if $Q<\infty$ and convex if $Q=\infty$ and $c=1$), we treat the two cases separately beginning with $Q<\infty$.

By defining $R_{n+1,\alpha}(i) \equiv V_{n,\alpha}(i) - V_{n,\alpha}(i+1)$, Equation 7 reveals the rather obvious fact that it is optimal to accept a customer of class x when n periods remain, the discount factor is $\alpha \geq 0$, and there are already i customers in the system if and only if r_x is at least as large as $R_{n,\alpha}(i)$. That is, given n, α, and i, there is a minimal reward that will be accepted. This is to be expected, for all customers classes have the same service time distribution. (See Cramer [7 , p.38-57] and Lippman and Ross [12] for two models in which service time depends upon the customer class.)

To garner more information about the behavior of $R_{n,\alpha}(i)$, the minimal acceptable reward, we need to know more about the behavior of $V_{n,\alpha}(i,x)$, the n period α-discounted return function. It is clear upon reflection that $V_{n,\alpha}(i,x)$ decreases in i and α and increases in n and x.

More important, however, is knowledge of

$$v_{n,\alpha}(i,x) \equiv V_{n,\alpha}(i,x) - V_{n,\alpha}(i+1,x)$$

(and, when $Q = \infty$, $v_{n,\alpha}(i) \equiv V_{n,\alpha}(i) - V_{n,\alpha}(i+1)$).

It will be shown that $v_{n,\alpha}(i,x)$ increases in i and in n and decrease in α, so that $R_{n,\alpha}(i)$ increases in i and in n and decreases in α. This knowledge will then be applied in seeking planning horizon results.

The next Theorem states the intuitively appealing idea that we become less eager to serve customers as the system fills up.

THEOREM 4: Given $\alpha \geq 0$, n, and x, the functions $V_{n,\alpha}(\cdot,x)$ and $V_{n,\alpha}(\cdot)$ are concave, so that $R_{n,\alpha}(\cdot)$ is a nondecreasing function.

Proof: Let $\alpha \geq 0$ be given. We claim that $H(i) \equiv (i \wedge c)V_n(i-1,0) + (c-(i \wedge c))V_n(i,0)$ is concave if $V_n(\cdot,0)$ is concave. To begin, let $f(j) = V_n(i+j,0)$ and take $i+2 \leq c$. Then

$$\begin{aligned} H(i) &- H(i+1) - [H(i+1) - H(i+2)] \\ &= i\, f(-1) + (c-i)f(0) - 2[(i+1)f(0) + (c-i-1)f(1)] \\ &\qquad + [(i+2)f(1) + (c-i-2)f(2)] \\ &= i\{[f(-1)-f(0)] - [f(0)-f(1)]\} + (c-i-2)\{[f(0)-f(1)] \\ &\qquad - [f(1)-f(2)]\}, \end{aligned}$$

and the concavity of $f(\cdot) = V_n(i+\cdot,0)$ implies that each term in braces is nonpositive. If $i+1=c$, the right side increases by $f(0) - f(-1) \leq 0$. If $i \geq c$, then the right side becomes $c\{[f(-1)-f(0)] - [f(0)-f(1)]\} \leq 0$, justifying the claim.

Next, we claim that concavity of $V_n(\cdot)$ implies that of $V_{n+1}(\cdot,x)$ for each x. To see this, fix x and let i_x be the smallest i for which $V_n(i) \geq r_x + V_n(i+1)$. To show that $f(i) \equiv v_{n+1}(i,x) - v_{n+1}(i+1,x) \leq 0$, we consider four cases. For $i<i_x-2$ and $i \geq i_x$ we have

$$f(i) = V_n(i+1) - V_n(i+2) - (V_n(i+2) - V_n(i+3))$$

and

$$f(i) = V_n(i) - V_n(i+1) - (V_n(i+1) - V_n(i+2))$$

respectively, and both are nonpositive by concavity of $V_n(\cdot)$. For $i=i_x-1$ and $i=i_x-2$ we have

$$f(i_x-1) = r_x - [V_n(i_x) - V_n(i_x+1)]$$

and

$$f(i_x-2) = V_n(i_x-1) - V_n(i_x) - [V_n(i_x) + r_x - V_n(i_x)]$$

respectively, and both are nonpositive by the definition of i_x. This justifies our claim that $V_{n+1}(\cdot,x)$ is concave if $V_n(\cdot)$ is concave.

Now $V_1(i,x) = r_x$ or 0 depending upon whether $i<c+Q$ or $i = c+Q$, so $V_1(\cdot,x)$ is concave. Coupling this fact with our first claim, we see that $V_1(\cdot)$ is concave, for the sum of concave functions is concave. Assume that $V_n(\cdot,x)$ is concave for each x. Then, as for *the case* $n=1$, it follows that $V_n(\cdot)$ is concave. But now our second claim yields the

desired result; namely, $V_{n+1}(\cdot,x)$ is concave for each x.

A simple consequence of the concavity of $V_{n-1}(\cdot)$ is the fact that $R_{n,\alpha}(i+1) \geq R_{n,\alpha}(i)$.

Q.E.D.

In addition to the decision maker's diminishing willingness to accept customers as the queue builds up (Theorem 4), our next result asserts that he becomes more and more selective as the length of the horizon increases. (This result also holds for the truly continuous time problem (see Theorem 7.3 of [17]).)

THEOREM 5: For each $\alpha \geq 0$, x, and i, $v_{n,\alpha}(i,x)$ is a nondecreasing function of n, so that $R_{n,\alpha}(i)$ is a nondecreasing function of n.

Proof: Setting $v_n(i) \equiv R_{n,\alpha}(i)$, we desire to show that

$$(S_n) \qquad v_n(i) \geq v_{n-1}(i) \quad , \qquad i=0,1,2,\cdots,c+Q,$$

holds for each $n \geq 1$. Since $v_1 \geq 0$ and $v_0 \equiv 0$, the statement S_1 is true.

Assume S_n is true. We claim that S_n implies C_{n+1} where C_{n+1} is defined by

$$(C_{n+1}) \qquad v_{n+1}(i,x) \geq v_n(i,x) \quad , \quad \text{all } i,x.$$

But C_{n+1} implies S_{n+1} as is easily seen from (8).

Thus, it only remains to show that S_n implies C_{n+1}. By Theorem 4, we need consider only six (instead of 16) cases. Fix x and define i_n

to be the smallest i for which $V_n(i) \geq r_x + V_n(i+1)$.

Case 1: $i+1<i_{n+1}$, $i+1<i_n$; by S_n we have

$$v_{n+1}(i,x) = v_n(i+1) \geq v_{n-1}(i+1) = v_n(i,x).$$

Case 2: $i+1 \geq i_{n+1}$, $i<i_{n+1}$, $i+1<i_n$

$$v_{n+1}(i,x) = r_x = r_x + V_{n-1}(i+1) - V_{n-1}(i+1) = V_n(i,x) - V_{n-1}(i+1) \geq v_n(i,x).$$

Case 3: $i \geq i_{n+1}$, $i+1<i_n$

$$v_{n+1}(i,x) \geq r_x = V_n(i,x) - V_{n-1}(i+1) \geq v_n(i,x).$$

Case 4: $i<i_{n+1} \wedge i_n \leq i+1$

$$v_{n+1}(i,x) = r_x = v_n(i,x).$$

Case 5: $i \geq i_{n+1}$, $i<i_n \leq i+1$

$$v_{n+1}(i,x) = v_n(i) \geq r_x = v_n(i,x).$$

Case 6: $i \geq i_{n+1} \wedge i_n$; by S_n we have

$$v_{n+1}(i,x) = v_n(i) \geq v_{n-1}(i) = v_n(i,x).$$

Q.E.D.

Monotone behavior of $R_{n,\alpha}(i)$ as a function of i and n was exhibited in Theorems 4 and 5, respectively. It is now shown that the minimal acceptable reward $R_{n,\alpha}(i)$ is a nonincreasing function of α. This is to be expected, for as α increases the future looks less attractive or promising whereas the present is not responsive to changes in α.

<u>THEOREM 6</u>: For each n, x and i, $v_{n,\alpha}(i,x)$ is a non-increasing function of α, *so that* $R_{n,\alpha}(i)$ is a nonincreasing function of α.

<u>Proof</u>: Writing $(\alpha+\Lambda)R_{n,\alpha}(i)$ as

$$\lambda \int_{\underline{X}} v_{n,\alpha}(i,y)p(dy) + (i\wedge c)\mu v_{n,\alpha}(i-1,0) + \mu[c-(i+1)\wedge c)]v_{n,\alpha}(i,0),$$

it is clear that $R_{n,\alpha}(i)$ is nonincreasing in α if for each x, $v_{n,\alpha}(i,x)$ is also.

Next, note $V_{1,\alpha}(i,x)$ and hence $v_{1,\alpha}(i,x)$ is constant in α. Now assume that $v_{n,\alpha}(i,x)$ is nonincreasing in α, so that the argument above shows that $R_{n,\alpha}(i)$ is nonincreasing in α. From Theorem 4 we know that $v_{n+1,\alpha}(i,x)$ assumes one of the following three values:

(i) $r_x + V_{n,\alpha}(i+1) - r_x - V_{n,\alpha}(i+2) = R_{n,\alpha}(i+1)$,

(ii) $r_x + V_{n,\alpha}(i+1) - V_{n,\alpha}(i+1) = r_x$,

(iii) $R_{n,\alpha}(i)$.

But each of these three expressions is nonincreasing in α.

Q.E.D.

From the definition of $R_{n,\alpha}(i)$, it is obvious that $R_{n,\alpha}(i) \leq r_K$, while Theorem 5 states that $R_{n,\alpha}(i)$ is nondecreasing in n, so that

$$R_\alpha(i) \equiv \lim_{n\to\infty} R_{n,\alpha}(i)$$

exists. Now if $r_x \geq R_\alpha(i)$ then $r_x \geq R_{n,\alpha}(i)$ for all n sufficiently

large, or if $r_x<R_\alpha(i)$ then $r_x<R_{n,\alpha}(i)$ for all n.
Furthermore, Theorems 4 and 5 reveal that $R_\alpha(i)$ is a nondecreasing function of i. Consequently, given $\alpha\geq 0$ and $x\epsilon\underline{\bar{x}}$, there are integers $i_{\alpha,x}$ and $N_{\alpha,x}<\infty$ such that

$$(10) \qquad r_x \begin{cases} \geq R_{n,\alpha}(i), & \text{if } i\leq i_{\alpha,x} \\ < R_{n,\alpha}(i), & \text{if } i>i_{\alpha,x} \end{cases} \qquad \text{whenever } n\geq N_{\alpha,x}.$$

Thus, as long as $N_{\alpha,x}$ or *more* periods remain, a class x customer can assert, without complete knowledge of n, that he will be accepted if $i\leq i_{\alpha,x}$ and rejected if $i>i_{\alpha,x}$.

This naturally raises the question of whether $N_{\alpha,x}$ will also suffice for all other customer classes in $\underline{\bar{x}}$. If the answer is affirmative, then $N_{\alpha,x}$ is called an α-planning horizon, and we write N_α instead of $N_{\alpha,x}$ to indicate that N_α works for all elements of $\underline{\bar{x}}$. If there is an integer N^* and an $\alpha^*>0$ such that N^* is an α-planning horizon for all $0<\alpha\leq\alpha^*$, then N^* is called a strong planning horizon.

While the existence of α-planning horizons is immediate if $\underline{\bar{x}}$ is a finite set, the situation is not so clear if $\underline{\bar{x}}$ is not finite nor is the existence of a strong planning horizon transparent even if $\underline{\bar{x}}$ is finite. The next several results reveal that α-planning horizons need not exist although weak α-planning horizons exist, that $R_\alpha(i)$ is a continuous function of α, and that there is a strong planning horizon if $\underline{\bar{x}}$ is finite.

EXAMPLE 1: An α-planning horizon need not exist.

Let p be Lebesgue measure on $\underline{X} = [1,2]$, $c=1$, $Q=0$, $r_x=x$, $\lambda=1$, $\mu = \frac{1}{4}$, and α rational.

A straightforward induction argument shows that $V_{n,\alpha}(i)$ is rational, $i=0,1$, α rational, and $n=0,1,2,\cdots$. (Of course $V_{n,\alpha}(i,x)$ is not necessarily rational.) Hence, $R_{n,\alpha}(0)$ is rational for all n and all rational α. Consequently, it suffices to show that $R_\alpha(0)$ is irrational.

Let $\alpha=0$ and define $g(\pi)$ to be the long-run expected return per unit time when we accept only those customers whose class is π or greater. Theorem 5 of Lippman and Ross [12] states that $g(\cdot)$ is unimodal while Theorem 3 of [12] establishes the optimality of this class of policies. For this example, we have

$$g(\pi) = \frac{2+\pi}{2} : (4 + \frac{1}{2-\pi}) = \frac{4-\pi^2}{2(9-4\pi)}, \quad 1 \leq \pi \leq 2.$$

Thus π^*, the optimal value of π, is $(9-\sqrt{17})/4$, an irrational number. But $\pi^* = R_0(0)$. This is seen as follows. Let V_n denote the expected n-period return obtained by accepting only customers of class π^* or higher. Then $V_n - n\Lambda g(\pi^*)$ is uniformly bounded (see Veinott [22, p. 1293]). Next, Theorem 5 can be employed to show that if $R_0(0) \neq \pi^*$, then there is an $\varepsilon > 0$ such that $|R_{n,0}(0)-\pi^*| > \varepsilon$ for all n sufficiently large, and from this it can then be shown that $V_{n,0} - n\Lambda g(\pi^*)$ goes to $-\infty$.

Q.E.D.

As evidenced in the example, the problem lies in the fact that if $<r_{x_m}>$ is a strictly increasing sequence with limit $R_\alpha(i)$ and if $R_{n,\alpha}(i)$ is strictly increasing, then $\{N_{\alpha,x_m}\}$ must be unbounded. We say that the finite integer W_α is a <u>weak α-planning horizon</u> if whenever $n \geq W_\alpha$ and the system is in state (i,x), the customer is accepted if $i \leq i_{\alpha,x}$ and rejected if $i > i_{\alpha,x}+1$, where $i_\alpha : \underline{x} \to \{0,1,\cdots,c+Q\}$ is defined in (10).

<u>THEOREM 7</u>: Assume that $\overline{\underline{x}}$ is closed, r is right-continuous, and let $\alpha \geq 0$ be given. Then there is a weak α-planning horizon and i_α is right-continuous.

<u>Proof</u>: Define $S_i = \{x : i_{\alpha,x} = i\}$ and $x_i = \inf\{x : x \in S_i\}$, $i=0,1,\cdots,c+Q$. We intend to show that $W_\alpha \equiv \max_i N_{\alpha,x_i}$ works.

Take $x, x' \in S_i$, $x \neq x_i \neq x'$ with $x' < x$. If $r_x = R_\alpha(i)$, then $r_{x'} < R_\alpha(i)$. But $x' \in S_i$, so we must have $r_x > R_\alpha(i)$. Since $\overline{\underline{x}}$ is closed, $x_i \in \overline{\underline{x}}$. Also, the right-continuity of r implies that $r_{x_i} \geq R_\alpha(i)$. But too, for $x \in S_i$ we have $r_{x_i} \leq r_x < R_\alpha(i+1)$. Hence, $x_i \in S_i$.

Q.E.D.

<u>LEMMA 2</u>: For each i and n, $V_{n,\alpha}(i)$ is a continuous function of α, so that $R_{n,\alpha}(i)$ is also continuous. Moreover, $R_\alpha(i)$ is a continuous function of α.

<u>Proof</u>: A straightforward induction argument establishes the continuity of $V_{n,\alpha}(i)$ and, hence, that of $R_{n,\alpha}(i)$.

Pick $\alpha' \geq 0$ and suppose $R_\alpha(i)$ is not continuous at α'. Then either $R_{\alpha'}(i) > R_\alpha(i) + \varepsilon$ for all $\alpha > \alpha'$ or $R_{\alpha'}(i) + \varepsilon < R_\alpha(i)$ for all $\alpha < \alpha'$, where $\varepsilon>0$ since $R_\alpha(i)$ is monotone. In the former case, this yields $R_{n,\alpha'}(i) > R_\alpha(i) + \frac{\varepsilon}{2} \geq R_{n,\alpha}(i) + \frac{\varepsilon}{2}$ for $\alpha>\alpha'$ and n large, contradicting the continuity of $R_{n,\alpha}(i)$ at α'. The latter case is similar.

Q.E.D.

The results of Lemma 2 and Theorems 4, 5, 6 are easily combined to yield

<u>THEOREM 8</u>: For each $x \in \bar{X}$ there are integers i_x $(=\lim_{\alpha\to 0^+} i_{\alpha,x})$ and M_x and a number $\alpha_x>0$ such that

(11)
$$r_x \begin{cases} \geq R_{n,\alpha}(i) \;, & i \leq i_x \\ < R_{n,\alpha}(i) \;, & i > i_x \end{cases} \qquad \text{whenever } n \geq M_x \text{ and } 0 \leq \alpha \leq \alpha_x.$$

In particular, there is a strong planning horizon if $\bar{X}$ is finite.

The analysis of the finite horizon problem, as embodied in Lemma 2 and Theorem 8, renders the infinite horizon problem, with or without discounting, practically trivial. In particular, we note the following consequences of Lemma 2 and Theorems 4-8: (1) if the system is in state (i,x), then it is optimal to accept the customer if and only if $r_x \geq R_\alpha(i)$, where $\alpha \geq 0$ is given. (2) The function $V_\alpha(\cdot)$ is concave and the return function $V_\alpha(\cdot,x)$ is concave and uniquely satisfies the functional equation of dynamic programming for $\alpha>0$. (3) Moreover, it is apparent from the continuity of $R_\alpha(i)$ that a strongly optimal policy will not, in general, exist if $\bar{X}$ is infinite; however, finiteness of $\bar{X}$ does ensure (via Theorem 8) the existence of a strongly optimal policy. (4) In the average cost problem, the functions $h(i,x) \equiv \lim_{\alpha\to 0^+} \{V_\alpha(i,x) - V_\alpha(0,0)\}$ and $h(i) \equiv \lim_{\alpha\to 0^+} \{V_\alpha(i) - V_\alpha(0)\}$ exist, are concave for each fixed x, satisfy the functional equation

(12) $$h(i,x) = \max\{r_x + h(i+1);\ h(i)\} - \overline{V}/\Lambda ,$$

and any stationary that selects an action maximizing the right side of (12) for each $s\varepsilon S$ is average optimal.

In addition, it is worth noting that, in the infinite horizon, the model is, in fact, the appropriate truly continuous time model and not merely some semi-Markov version. That this is so is a result of the fact that there are no actions available unless an arrival occurs. This is in contrast to the model of section 3, for there one can turn servers on or off at any time; but of course it can be shown that to do so at any time other than an arrival or departure is suboptimal [23]. On the other hand, it should be clear that the model with n transitions, as embodied in Equations 7 and 8, is not equivalent to the truly continuous time model with time horizon n/Λ. Finally, we note that the functional equations (12) and (7) with $n=+\infty$ can be rewritten so that the fictitious events are eliminated.

We conclude by establishing that 0-optimal policies do exist even if $\underline{X}$ is infinite and by providing a partial ordering on the (nonempty) set of 0-optimal policies. Utilizing the characterization inherent in this partial ordering, we then show that the maximally accepting policy among the set of policies satisfying the functional equation (12) is 0-optimal. In particular, if $\underline{X}$ is finite, then this maximally accepting policy is, in fact, strongly optimal, so that the further computation always associated with finding strongly optimal (and 0-optimal) policies need not be performed (see Miller and Veinott [18a] for the usual algorithm).

THEOREM 9. The stationary policy R, defined by

$$R(i) = \lim_{\alpha \to 0^+} R_\alpha(i) ,$$

exists and is 0-optimal. Moreover, suppose two stationary policies π and σ are 0-optimal and that there is a state (i',x') such that $\pi(i',x) = 1 \neq \sigma(i',x')$ and $\pi(s) = \sigma(s)$ for $s \neq (i',x')$. Then for each $\alpha > 0$, the α-discounted return of π exceeds that of σ. In particular, if there is a strongly optimal stationary policy, then it can be characterized as the maximally accepting policy among the class of stationary 0-optimal policies.

Proof: First, reformulate the problem so that the state space is finite and the action space is uncountable. This is accomplished by letting the state be the number of customers already accepted into the queue and the action space be the set of subsets of $\underline{X}$. Here, each action specifies the set of customer classes that will be admitted into the queue. It is evident from Theorem 4 that the action space can be further reduced to those subsets of the form $(a,K] \cap \underline{X}$ or $[a,K] \cap \underline{X}$.

The stationary policy $R(i) = \lim_{\alpha \to 0^+} R_\alpha(i)$ exists as $\underline{X}$ is closed by hypothesis and $R_\alpha(i)$ is a monotone function of α by Theorem 6. Using Blackwell's representation for the return $V_\beta(\pi)$ of a stationary policy π [3a], we have

$$V_\beta(\pi) = \frac{x_\pi}{1-\beta} + y_\pi + \varepsilon(\beta,\pi) \ ,$$

where $\beta = \Lambda/(\alpha+\Lambda)$. The vectors x_π and y_π are the unique solutions of

$$x = Q_\pi^* r_\pi$$

and

$$(I-Q_\pi)y = r_\pi - x_\pi, \quad Q_\pi^* y = 0 \ ,$$

whereas the vector function $\varepsilon(\beta,\pi)$ is given by

$$\varepsilon(\beta,\pi) = (H(\beta,\pi)-H(\pi))r_\pi \ .$$

In the above Q_π is the one-step transition matrix associated with policy π, Q_π^* is the stationary distribution, and $r_\pi(i)$ is the expected immediate reward when action $\pi(i)$ is chosen while in state i. Finally,

$$H(\beta,\pi) = [I - \beta(Q_\pi - Q_\pi^*)]^{-1} - Q_\pi^*$$

and

$$H_\pi = [I - Q_\pi + Q_\pi^*]^{-1} - Q_\pi^* \ .$$

It is evident that the convergence of R_α to R implies that of Q_{R_α} to Q_R, $Q_{R_\alpha}^*$ to Q_R^*, and r_{R_α} to r_R. Consequently, $x_{R_\alpha} \nearrow x_R$ and $y_{R_\alpha} \to y_R$. Hence, to show R is 0-optimal, it suffices to show that

$$\varepsilon(\beta,R) - \varepsilon(\beta,R_\beta) \to 0 \quad \text{as} \quad \alpha \to 0 \ .$$

Because $I-\beta(Q_{R_\alpha}-Q_{R_\alpha}^*)$ and $I-\beta(Q_R-Q_R^*)$ are nonsingular for $0 < \beta \leq 1$, the convergence of Q_{R_α}, $Q^*_{R_\alpha}$ and r_{R_α} yields $\varepsilon(\beta,R)-\varepsilon(\beta,R_\beta) \to 0$ as $\alpha \to 0$.

To show that the α-discounted return of π exceeds that of σ, observe that, by hypothesis

$$r_{x'} = \lim_{\alpha\to 0^+} \{V_\alpha(i') - V_\alpha(i'+1)\} ,$$

so by Theorem 6 we have

$$r_{x'} > V_\alpha(i') - V_\alpha(i'+1), \text{ all } \alpha > 0 .$$

Q.E.D.

Utilizing Theorem 9, we can now present an efficient algorithm for computing the 0-optimal policy R which is a strongly optimal policy if one exists as is true whenever $\underline{X}$ is finite. First, however, we need the following result concerning Markov-decision processes.

LEMMA 3. Consider a Markov decision process with state space S, action space A, and bounded reward function. Denote the β-discounted optimal return function by V_β, the β-discounted return function of the stationary policy π by $V_{\beta,\pi}$, and their difference by $\varepsilon(\beta,\cdot)$. Then if the stationary policy π which selects action $\pi(s)$ from state s is 1-optimal and if $\varepsilon(\beta,s)$ is uniformly bounded, $\pi(s)$ is a maximizer of the function equation

$$h(s) = \max_{a\in A} \{r(s,a) + \int_S h(s')dP_{s,s'}(a)\} - \overline{V} .$$

Proof: We have

$$V_\beta(s) = \max_{a\varepsilon A} \{r(s,a) + \beta \int_S V_\beta(s')dP_{s,s'}(a)\}$$

$$\leq r(s,\pi_s) + \beta \int_S V_{\beta,\pi}(s')dP_{s,s'}(\pi_s) + \varepsilon(\beta,s)$$

$$= r(s,\pi_s) + \beta \int_S V_\beta(s')dP_{s,s'}(\pi_{s'}) + \varepsilon(\beta,s) - \beta \int_S (V_\beta(s')-V_{\beta,\pi}(s'))dP_{s,s'}(\pi_s) \; ,$$

where $\varepsilon(\beta,s) \equiv V_\beta(s)-V_{\beta,\pi}(s)$. Since π is 1-optimal, $\varepsilon(\beta,s) \to 0$ as $\beta \to 1$ for each s. Furthermore, since $\int_S dP_{s,s'}(\pi_s)$ is finite $(= 1)$ and $\varepsilon(\beta,s)$ is uniformly bounded, $\int_S \varepsilon(\beta,s')dP_{s,s'}(\pi_s) \to 0$ as $\beta \to 1$ for each s. Therefore, defining $h(s) = \lim_{\beta\to 1^-}[V_\beta(s)-V_\beta(0)]$, we obtain

$$h(s) \leq r(s,\pi_s) + \lim_{\beta\to 1^-}\{\beta \int_S (V_\beta(s')-V_\beta(0)]dP_{s,s'}(\pi_s) - (1-\beta)V_\beta(0)$$

$$+ \varepsilon(\beta,s) + \int_S \varepsilon(\beta,s')dP_{s,s'}(\pi_s)\}$$

$$= r(s,\pi_s) + \int_S h(s')dP_{s,s'}(\pi_s) - \overline{V} \leq \max_{a\varepsilon A} \{r(s,a) + \int_S h(s')dP_{s,s'}(a)\} - \overline{V}$$

$$= h(s) \; .$$

Q.E.D.

<u>COROLLARY 1</u>. The stationary policy R is the maximally accepting policy that satisfies the functional equation (12); that is, accepting whenever $r_x + h(i+1) \geq h(i)$ (where h is any solution of (12)) is 0-optimal and strongly optimal if there is a strongly optimal policy.

<u>Proof</u>: By Lemma 3, every 0-optimal policy including R satisfies (12), whereas Theorem 9 shows that R is maximally accepting among the class of 0-optimal policies.

Let π be any policy satisfying the functional and suppose that $\pi(i',x') = 1$, yet $R(i') > r_{x'}$; then $r_{x'} + h(i'+1) = h(i')$.

Since any h that satisfies (12) can be written as

$$h(i,x) = \lim_{\alpha\to 0^+}[V_\alpha(i,x)-V_\alpha(0,0)] + c ,$$

where c is some finite constant independent of i and x, it follows from (12) that

$$r_{x'} = \lim_{\alpha\to 0^+}[V_\alpha(i')-V_\alpha(i'+1)] = \lim_{\alpha\to 0^+} R_\alpha(i') = R(i') .$$

This contradicts the fact that $R(i') > r_{x'}$.

Q.E.D.

THE CASE $Q = \infty$

When the queue capacity Q is infinite, it is imperative that r_α be given by (9), for otherwise the optimal decision is always accept the customer requesting admission. Unfortunately, utilizing (9) renders this case inherently more complicated than the case $Q<\infty$. For example, the function $V_{1,\alpha}(\cdot,x)$ is not concave, nor is it convex unless $c=1$. Consequently, we limit consideration to the case $c=1$. And although $v_{n,\alpha}(\cdot,x)$ and $v_{n,\alpha}(\cdot)$ are convex, this is not sufficient to yield the analog of Theorem 4. The final result of Theorem 4, however, is rather easily obtained as are the results of Theorem 5. But the results of an appropriately modified analog of Theorem 6 do not appear to hold, so investigation of the sensitivity of the solution as a function of α is, at present, not possible.

We begin by noting that the minimal acceptable reward $R_{n,\alpha}(i)$ is not $v_{n,\alpha}(i)$ but rather $v_{n,\alpha}(i)/\beta^i$, where $\beta \equiv \mu/(\alpha+\mu)$. This can be seen from inspection of the recursive equations ($\delta_i=1$ for $i>0$ and 0 for $i=0$; $\Lambda=\lambda+\mu$)

(13) $$V_{n+1,\alpha}(i,x) = \max\{r_x\beta^i + V_{n,\alpha}(i+1);\ V_{n,\alpha}(i)\},\quad x\in\underline{X}\cup\{0\},\ i=0,1,\cdots$$

and

(14) $$V_{n,\alpha}(i) = \frac{1}{\alpha+\Lambda}\{\lambda\int_{\underline{X}} V_{n,\alpha}(i,y)p(dy) + \delta_i\mu V_{n,\alpha}(i-1,0)\}.$$

As in the case $Q<\infty$, it remains true that the decision maker becomes more discriminating both as the horizon lengthens and as the system fills up.

<u>THEOREM 10</u>: Given $\alpha\geq 0$, n, and x, the functions $v_{n,\alpha}(i,x)/\beta^i$ and $R_{n,\alpha}(i)$ are both nondecreasing in i and nondecreasing in n.

<u>Proof</u>: Since $v_1(i,x)/\beta^i = r_x(1-\beta)$ and $v_1(i)/\beta^i = \lambda(1-\beta)\int_{\underline{X}} r_y p(dy) = v_1(i+1)/\beta^{i+1}$, the result is true for $n=1$. Assume that $v_n(i,x) \leq v_n(i+1,x)/\beta$, all x, all i. Then

$$v_n(i) = \frac{1}{\alpha+\Lambda}\{\lambda\int_{\underline{X}} v_n(i,x)p(dx) + \mu\delta_i v_n(i-1,0)\}$$

$$\leq \frac{1}{\beta}\frac{1}{\alpha+\Lambda}\{\lambda\int_{\underline{X}} v_n(i+1,x)p(dx) + \mu v_n(i,0)\} = \frac{1}{\beta}v_n(i+1).$$

Using $v_n(i) \leq v_n(i+1)/\beta$, we only need consider four cases to complete the induction argument.

Case 1: Accept at i, $i+1$, $i+2$.

$$v_{n+1}(i,x) = r_x\beta^i(1-\beta) + v_n(i+1) \leq r_x\beta^i(1-\beta) + v_n(i+2)/\beta = v_{n+1}(i+1)/\beta.$$

Case 2: Accept at i and $i+1$.

$$v_{n+1}(i,x) \leq r_x\beta^i = v_{n+1}(i+1,x)/\beta.$$

Case 3: Accept at i.

$$v_{n+1}(i,x) = r_x\beta^i \leq v_n(i+1)/\beta = v_{n+1}(i+1)/\beta.$$

Case 4: Do not accept at i, $i+1$, $i+2$.

$$v_{n+1}(i,x) = v_n(i) \leq v_n(i+1)/\beta = v_{n+1}(i+1,x)/\beta.$$

The proof that the two functions do not decrease as n increases is nearly the same as the proof of Theorem 5.

Q.E.D.

V. AN M/M/1 QUEUE WITH VARIABLE SERVICE RATE

The model considered in this section is a generalization of a model first considered by Crabill [5,6] and later modified by Sabeti [20]. Here, we have an M/M/1 system with arrival rate λ, service rate μ, and infinite queue capacity. The decision variable is the exponential service rate μ to be employed, where μ lies in some subset A of $[0, \bar{\mu}]$, $\bar{\mu}<\infty$. The cost structure consists of three parts: a holding cost h per customer per unit time, a service cost c_μ per unit time when the service rate is μ, and a reward of $R\geq 0$ that is received whenever a customer completes service. Of course, c is taken to be a strictly increasing function with $c_0\geq 0$ and $c_{\bar{\mu}}<\infty$. To ensure the existence of an α-optimal n period policy, we assume that c is left continuous and that A is a closed set.

Here, the state of the system is simply the number of customers in the system, so the state space S is $\{0,1,2,\cdots\}$ and the action space is the set A, previously defined. Taking $\Lambda = \lambda + \bar{\mu}$, the law of motion q is given by $(q(-1|0,\mu) = 0$ and $q(0|0,\mu) = \bar{\mu}/\Lambda)$

$$q(i+1|i,\mu) = \frac{\lambda}{\Lambda},\quad q(i|i,\mu) = \frac{\bar{\mu}-\mu}{\Lambda} \text{ and } q(i-1|i,\mu)= \frac{\mu}{\Lambda},$$

while $r_\alpha(i,\mu) = (c_\mu + hi - \mu R)/(\alpha+\Lambda)$. This easily leads to the recursive equations

$$V_{n+1,\alpha}(i) = \frac{1}{\alpha+\Lambda} \min_{\mu\varepsilon A} \{c_\mu + hi + V_{n,\alpha}(i,\mu)\}, \tag{15}$$

where $(V_{n,\alpha}(0,\mu) = V_{n,\alpha}(0,0)$, all $\mu \varepsilon A)$

(16) $$V_{n,\alpha}(i,\mu) = \lambda V_{n,\alpha}(i+1) + \bar{\mu} V_{n,\alpha}(i) - \mu(V_{n,\alpha}(i) - V_{n,\alpha}(i-1)) - \mu R.$$

The model, as incorporated in Equations (15) and (16), generalizes Crabill's model in that A is not required to be a finite set and R is not required to be 0. On the other hand, Crabill lets h_i be the holding cost rate where $h_.$ is an unbounded nondecreasing function, so the case $h_i = h \cdot i$ is included by Crabill. We can, however, relax the assumption that $h_i = h \cdot i$ and assume only that h_i is an unbounded nondecreasing convex function. This relaxation is also possible for the model of section 3; and the proofs in both sections 3 and 5 go through without change. We shall consider finite and infinite horizon problems with and without discounting, whereas Crabill restricted his investigation to the class of stationary policies for the infinite horizon average expected cost case.

Sabetti's model assumed A finite, $h=0$, and a finite queue capacity, again with average cost as the criterion of optimality. As the changes needed to incorporate this model are fairly straightforward, we concentrate solely upon the model as posed in Equations 15 and 16.

To begin, define $\mu^*_{n,\alpha}(i)$ to be an optimal service rate when the system is in state i, n periods remain and the discount factor is $\alpha \geq 0$. Also, define

$$v_{n,\alpha}(i) = V_{n,\alpha}(i) - V_{n,\alpha}(i-1),$$

$$v_{n,\alpha}(i,\mu) = V_{n,\alpha}(i,\mu) - V_{n,\alpha}(i-1,\mu)$$

$$D_{n,\alpha}(i) = \min_{\mu} \{c_\mu - \mu(v_{n,\alpha}(i) + R)\} - \min_{\mu} \{c_\mu - \mu(v_{n,\alpha}(i-1) + R)\}.$$

As in section IV, the behavior of μ^* will be gleaned from that of these three functions.

A policy possessing the intuitively appealing property that more customers in the system leads to a faster service rate (i.e., $\mu^*_{n,\alpha}(i)$ is nondecreasing as a function of i) is termed a <u>connected</u> or <u>switch-over</u> policy. Crabill's principal result states that if attention is restricted to the class of stationary policies, then there is a switch-over policy that is average optimal. In showing that $V_{n,\alpha}(\cdot)$ is convex, we extend this result to the finite horizon problems and to the infinite horizon discounted problem, without restriction to the class of stationary policies. From this result, we proceed by showing that $\mu^*_{n,\alpha}(i)$ increases with n and decreases with α so that there is a strong planning horizon and, consequently, a strongly optimal policy if A is finite.

Before presenting our first result, it should be noted that although the introduction of the reward R introduces the possibility of $V_{n,\alpha}(i)$ being negative, a straightforward induction argument establishes the hoped for fact that $V_{n,\alpha}(i)$ is a nondecreasing function of i; i.e., $v_{n,\alpha}(i) \geq 0$.

<u>THEOREM 11</u>: For each $\alpha \geq 0$ and n, $V_{n,\alpha}(\cdot)$ is convex, so that $\mu^*_{n,\alpha}(i)$ is a nondecreasing function *of i*; that is, $\mu^*_{n,\alpha}$ is a switch-over policy. In addition $V_{n,\alpha}(\cdot,\mu)$ is convex for each $\mu \in A$.

Proof: First fix $\alpha \geq 0$ and note that $V_1(\cdot)$ is linear, and assume that $V_n(\cdot)$ is convex. Then setting $\mu^*_{n+1,\alpha}(i) = \mu_i$ we have

$$(\alpha+\Lambda)v_{n+1}(i+1) \geq c_{\mu_{i+1}} - c_{\mu_{i+1}} + \lambda v_n(i+2) + \bar{\mu}\, v_n(i+1) - \mu_{i+1}[v_n(i+1)-v_n(i)].$$

Similarly,

$$(\alpha+\Lambda)v_{n+1}(i) \leq c_{\mu_{i-1}} - c_{\mu_{i-1}} + \lambda v_n(i+1) + \bar{\mu}\, v_n(i) - \mu_{i-1}[v_n(i)-v_n(i-1)].$$

Combining these two inequalities, we obtain

$$v_{n+1}(i+1) - v_{n+1}(i) \geq \lambda[v_n(i+2)-v_n(i+1)] + (\bar{\mu}-\mu_{i+1})[v_n(i+1)-v_n(i)]$$
$$+ \mu_{i-1}[v_n(i)-v_n(i-1)] \geq 0$$

since $\mu_{i+1} \leq \bar{\mu}$, $\mu_{i-1} \geq 0$, and $V_n(\cdot)$ is convex by hypothesis. This completes the induction argument.

Next, observe that $V_{n+1}(\cdot)$ can be written as follows:

(17) $$V_{n+1,\alpha}(i) = \frac{1}{\alpha+\Lambda}\{hi + \lambda V_{n,\alpha}(i+1) + \bar{\mu}\, V_n(i) + \min_{\mu} [c_\mu - \mu(v_{n,\alpha}(i)+R)]\}.$$

The existence of a μ in $\bar{X}$ minimizing $c_\mu - \mu(v_{n,\alpha}(i)+R)$ follows from c being left continuous and $\bar{X}$ closed. Since $V_{n,\alpha}(\cdot)$ is convex, $v_{n,\alpha}(i+1) \geq v_{n,\alpha}(i)$ so the desired result -- namely $\mu^*_{n+1,\alpha}(i+1) \geq \mu^*_{n+1,\alpha}(i)$ -- follows immediately when this last fact is coupled with c_μ strictly increasing and (17).

Q.E.D.

REMARK. A closer inspection of Equation 17 reveals that if c_μ is continuous, then $\mu^*_{n,\alpha}(i+1) \geq \mu^*_{n,\alpha}(i)$ even if c_μ is not a nondecreasing function.

Using the usual definition of a transition, Prabhu and Stidham have obtained Theorem 11 by appending the following assumption: c_μ is itself a convex function of μ.

THEOREM 12: For each $\alpha \geq 0$ and $i \in S$, $v_{n,\alpha}(i)$ is a nondecreasing function of n, so that $\mu^*_{n,\alpha}(i)$ is also a nondecreasing function of n.

Proof: From (17), it is clear that the desired monotonicity of $\mu^*_{n,\alpha}(i)$ is a simple consequence of $v_{n+1,\alpha}(i) \geq v_{n,\alpha}(i)$. Since $v_0(i) \equiv 0$ and $v_1(i) \geq 0$, the result is true for $n=0$. Assume it true for $n-1$; i.e., $v_{n,\alpha}(i) \geq v_{n-1,\alpha}(i)$. Utilizing (17), it can easily be seen that $v_{n+1,\alpha}(i) \geq v_{n,\alpha}(i)$ if $\bar{\mu}\, v_{n,\alpha}(i) + D_{n,\alpha}(i) \geq \bar{\mu}\, v_{n-1,\alpha}(i) + D_{n-1,\alpha}(i)$. Writing $\mu^*_{n,\alpha}(i) = \mu(n,i)$, we have

$$D_{n,\alpha}(i) \geq c_{\mu(n+1,i)} - \mu(n+1,i)[v_n(i)+R] - \{c_{\mu(n,i-1)} - \mu(n,i-1)[v_n(i-1)+R]\}$$

and

$$D_{n-1,\alpha}(i) \leq c_{\mu(n+1,i)} - \mu(n+1,i)[v_{n-1}(i)+R] - \{c_{\mu(n,i-1)} - \mu(n,i-1)[v_{n-1}(i-1)+R]\}$$

and thus

$$\bar{\mu}\, v_n(i) + D_n(i) - [\bar{\mu}\, v_{n-1}(i) + D_{n-1}(i)]$$

$$\geq [\bar{\mu}-\mu(n+1,i)][v_n(i) - v_{n-1}(i)] + \mu(n,i-1)[v_n(i-1)-v_{n-1}(i-1)] \geq 0.$$

Q.E.D.

The existence of the limit on n of $\mu^*_{n,\alpha}(i)$ is immediate from Theorem 12, and we would expect this limit to be α-optimal provided that $\lim_{n\to\infty} V_{n,\alpha}(i)$ exists. Clearly $\langle V_{n,\alpha}(i)\rangle_{n=1}^{\infty}$ is a bounded sequence (see proof of Theorem 2 for the necessary technique), but is it monotone? Evidently not if $R>0$. Nevertheless, the limit does exist, and we have

THEOREM 13: For $\alpha>0$, $V_\alpha(\cdot)$ exists, is convex, and is the unique solution to the functional equation of dynamic programming. Moreover, the stationary policy μ_α^*, defined by

$$\mu_\alpha^*(i) = \lim_{n\to\infty} \mu_{n,\alpha}^*(i),$$

is α-optimal and is a switch-over policy.

Proof: If rate $\mu_{m+j,\alpha}^*(i)$ is employed rather than $\mu_{j,\alpha}^*(i)$ when j periods remain, $j=1,2,\cdots,n$, then an upper bound on $V_{n,\alpha}(i)$ is obtained and the first n periods contribute no difference in the cost between $V_{n+m}(i)$ and the bound for $V_n(i)$, and we obtain

$$V_{n+m,\alpha}(i) - V_{n,\alpha}(i) \geq \left(\frac{\Lambda}{\alpha+\Lambda}\right)^n \left(-\frac{\bar{\mu} R}{\alpha+\Lambda}\right) \sum_{j=0}^{m-1} \left(\frac{\Lambda}{\alpha+\Lambda}\right)^j \geq \left(\frac{\Lambda}{\alpha+\Lambda}\right)^n (-\bar{\mu}R/\alpha).$$

Similarly, we obtain

$$V_{n+m,\alpha}(i) - V_{n,\alpha}(i) \leq \left(\frac{\Lambda}{\alpha+\Lambda}\right)^n \sum_{j=0}^{\infty} [c_{\bar{\mu}} + (i+n+j)h]\left(\frac{\Lambda}{\alpha+\Lambda}\right)^j < \infty.$$

Consequently, $\langle V_{n,\alpha}(i)\rangle_{n=1}^{\infty}$ is a Cauchy sequence for each i and each $\alpha>0$, and thus $V_\alpha(\cdot)$ exists.

The remaining facts follow as in the proof of Theorem 2 except for the α-optimality of μ_α^* which (see Equation 17) follows from $\mu_{n,\alpha}^*(i)$ being a minimizer of $c_\mu-\mu(v_{n,\alpha}(i)+R)$ and $\langle v_{n,\alpha}(i)\rangle_{n=1}^{\infty}$ being a non-decreasing sequence with limit $v_\alpha(i)$. That μ_α^* is a switch-over policy is a consequence of Theorem 11.

Q.E.D.

THEOREM 14: For each $n \geq 1$ and i, $v_{n,\alpha}(i)$ is a strictly decreasing function of α, so that $\mu^*_{n,\alpha}(i)$ is a nonincreasing function of α.

Proof: The result holds for $n=1$ as $v_{1,\alpha}(i) = h/(\alpha+\Lambda)$. Assume it holds for n. Taking $\alpha_1<\alpha_2$, writing $\mu^*_{n+1,\alpha}(i) = \mu(\alpha,i)$, and employing (17), we have, as in the proof of Theorem 12,

$$(\alpha_1+\Lambda)v_{n+1,\alpha_1}(i) - (\alpha_2+\Lambda)v_{n+1,\alpha_2}(i)$$
$$\geq [\bar{\mu}-\mu(\alpha_1,i)][v_{n,\alpha_1}(i)-v_{n,\alpha_2}(i)] + \mu(\alpha_2,i-1)[v_{n,\alpha_1}(i-1)-v_{n,\alpha_2}(i-1)] \geq 0,$$

so $v_{n+1,\alpha_1}(i) \geq v_{n+1,\alpha_2}(i)\,(\alpha_2+\Lambda)/(\alpha_1+\Lambda) > v_{n+1,\alpha_2}(i)$.

Q.E.D.

In general, the existence of strongly optimal policies necessitates a finite state and action space (see Denardo [8, p. 487]). Consequently, it came as no surprise that in order to ensure the existence of a strong planning horizon in the optimal customer selection model, it was necessary to assume $\underline{x}$ finite (thereby rendering a problem with finite state and action space). In the model with variable service rate, however, the problem of an infinite state space cannot be assumed away. Fortunately, the next Lemma establishes that the model can essentially be reduced to one with a finite set of states -- at least for n large and α small -- in that $\bar{\mu}$ is the optimal rate for all large states.

LEMMA 4: If $\bar{\mu}$ is an isolated point of A or if $\{(c_{\bar{\mu}}-c_\mu)/(\bar{\mu}-\mu):\mu\varepsilon A\sim\{\bar{\mu}\}\}$ is bounded, then there are numbers $N^*<\infty$, $i^*<\infty$ and $\alpha^*>0$ such that

$\mu^*_{n,\alpha}(i) = \bar{\mu}$ whenever $n \geq N^*$, $i \geq i^*$, and $0 \leq \alpha \leq \alpha^*$.

Proof: We claim that[1]

(18) $$V_{n,\alpha}(i+1) \geq \frac{h}{\alpha+\Lambda} \sum_{j=0}^{n \wedge i} \left(\frac{\Lambda}{\alpha+\Lambda}\right)^j .$$

This is seen as follows. First,

$$V_{n+1,\alpha}(i) \geq \frac{1}{\alpha+\Lambda} \{h+\lambda V_{n,\alpha}(i+1)+\mu^*_{n+1,\alpha}(i)V_{n,\alpha}(i-1)+(\bar{\mu}-\mu^*_{n+1,\alpha}(i))V_{n,\alpha}(i)\}$$

$$\geq \frac{1}{\alpha+\Lambda} \{h+ V_{n,\alpha}(i-1)\},$$

where the last inequality results from the convexity of $V_{n,\alpha}(\cdot)$. Iterating this inequality yields (18) as claimed.

Define $\mathcal{V}_{n,\alpha}(i)$ to be the return of the strategy $\sigma_{n,\alpha} = \langle\sigma_{n,\alpha,m}(\cdot)\rangle$ with service rates given by

$$\sigma_{n,\alpha,m}(i) = \begin{cases} \bar{\mu} & , \quad \text{if } m=n \\ \mu^*_{m,\alpha}(i) , & \text{if } m<n. \end{cases}$$

Then

(19) $$V_{n,\alpha}(i)-\mathcal{V}_{n,\alpha}(i) \geq \frac{1}{\alpha+\Lambda} \{c_{\mu^*_{n,\alpha}(i)}-c_{\bar{\mu}}+(\bar{\mu}-\mu^*_{n,\alpha}(i))V_{n-1,\alpha}(i)\}.$$

Let $\varepsilon = \bar{\mu} - \sup\{\mu^*_{n,\alpha}(i): n \geq 1,\ \alpha>0,\ i \in S\}$. If $\varepsilon>0$, then it follows from (18) and (19) that $V_{n,\alpha}(i)-\mathcal{V}_{n,\alpha}(i)$ is strictly positive for some finite i,

---

[1] The bound given in Equation 18 suffices to correct an error in Equation 31 of reference [11].

and n and strictly positive α, a contradiction.

Consequently, let us suppose that $\varepsilon=0$. If $\mu^*_{n,\alpha}(i) = \bar{\mu}$ some n,i and $\alpha>0$, then the desired result follows from Theorems 11, 12, and 14. Therefore, assume that $\bar{\mu}$ is not an isolated point of A and that there is a set $\{\mu^*_{n,\alpha}(i)\}$ with supremum $\bar{\mu}$ and $\bar{\mu}$ not in this set. By Theorems 11, 12, and 14, there is a sequence $<\mu^*_{n,\alpha_n}(i_n)>$ from this set with $0<\alpha_{n+1}<\alpha_n$, $i_{n+1}>i_n$ and limit $\bar{\mu}$. Since $(c_{\mu^*_{n,\alpha_n}}(i_n)-c_{\bar{\mu}})/(\bar{\mu}-\mu^*_{n,\alpha_n}(i_n))$ is bounded below by hypothesis and our bound on $<v_{n-1,\alpha_n}(i_n)>$ is non-decreasing with limit $+\infty$ by (18), (19) reveals that $V_{n,\alpha_n}(i_n)-V_{n,\alpha_n}(i_n)$ is strictly positive for some finite n.

Q.E.D.

Theorems 11, 12, 14 and Lemma 4 yield the following strong planning horizon theorem where μ^* is defined by

$$\mu^*(i) = \lim_{\alpha\to 0^+} \mu^*_\alpha(i) .$$

<u>THEOREM 15</u>: If A is finite, then there is an $N^*<\infty$ and $\alpha^* 0$ such that for each i,

$$\mu^*_{n,\alpha}(i) = \mu^*(i), \quad \text{whenever} \quad n\geq N^* \quad \text{and} \quad 0<\alpha\leq\alpha^*.$$

In particular, μ^* is strongly optimal.

In Example 2 below, it is demonstrated that a strong planning horizon may exist even if A is infinite, but there is no reason even to believe that a strongly optimal policy will always exist. However, we make the following conjecture: if c has a bounded derivative, then the stationary policy μ^* defined by $\mu^*(i) = \lim_{\alpha\to 0^+} \mu_\alpha(i)$ is 0-optimal.

EXAMPLE 2: A strong planning horizon may exist even if A is infinite. Suppose $A = [0,\bar{\mu}]$ and c is linear; that is,

$$c_\mu = K\mu, \quad K > 0 .$$

Then $\mu^*_{n,}(i)=0$ or $\bar{\mu}$ depending whether $K \geq v_{n-1,\alpha}(i)+R$ or $K<v_{n-1,\alpha}(i)+R$.

Let $H(i) = \lim_{\substack{\alpha\to 0^+ \\ n\to\infty}} v_{n,\alpha}(i)+R$ and define μ^* by

$$\mu^*(i) = \begin{cases} 0, & \text{if } K \geq H(i) \\ \bar{\mu}, & \text{if } K<H(i). \end{cases}$$

Then it is clear from Theorems 11, 12, and 14 that μ^* is strongly optimal and that there is a strong planning horizon. (Notice that the "bang-bang" form of μ^* does not depend upon linearity in the holding cost function.)

Q.E.D.

THEOREM 16: If $\lambda<\bar{\mu}$ and if either $\bar{\mu}$ is an isolated point of A or $\{c_{\bar{\mu}}-c_\mu)/(\bar{\mu}-\mu); \mu\varepsilon A\setminus\{\bar{\mu}\}\}$ is a bounded set, then h, defined by $h(i) = \lim_{\alpha\to 0^+} [V_\alpha(i)-V_\alpha(0)]$, is convex and satisfies the functional equation

(20) $$h(i) = \frac{1}{\Lambda}\{h\cdot i+\lambda h(i+1)+\bar{\mu}h(i) + \min_{\mu\varepsilon A}[c_\mu-\mu(h(i)-h(i-1)+R)]\} - \bar{V}/\Lambda ,$$

and any stationary policy that selects a rate which minimizes the right side of (20) for each i is average optimal. In particular, μ^* is average optimal.

Proof: By Lemma 4 and $\lambda<\bar{\mu}$, the assumptions of Theorem 4 of [11] are satisfied so that $h(\cdot)$ exists, satisfies (20), and any stationary policy that selects a rate which minimizes the right side of (20) is average optimal, while convexity follows from Theorem 13. Left continuity of

c_μ and $\mu_\alpha^*(i)$ nonincreasing in α yield

$$\min_{\mu\varepsilon A}\{c_\mu-\mu(R+h(i)-h(i-1))\}=\lim_{\alpha\to 0^+}\min_{\mu\varepsilon A}\{c_\mu-\mu[R+V_\alpha(i)-V_\alpha(i-1)]\},$$

$$=\lim_{\alpha\to 0^+}\{c_{\mu_\alpha^*(i)}-\mu_\alpha^*(i)[R+V_\alpha(i)-V_\alpha(i-1)]\}$$

$$=c_{\mu^*(i)}-\mu^*(i)[R+h(i)-h(i-1)],$$

so that μ^* is, indeed, average optimal.

Q.E.D.

In the model of optimal customer selection, we were able to characterize a strongly optimal policy (if it exists) as the maximally accepting policy in the class of 0-optimal policies. For the variable service model the characterization asserts that, given a choice, slower rates are preferred.

THEOREM 17: Suppose two stationary policies π and σ are 0-optimal and that $\pi(i) = \sigma(i)$ for $i\neq i'$ and $\pi(i') = \hat{\mu}<\tilde{\mu} = \sigma(i')$. Then for each $\alpha>0$, the α-discounted cost of π is less than that of σ.

Proof: By hypothesis, $\hat{\mu}<\tilde{\mu}$ and

$$\lim_{\alpha\to 0^+}[c_{\hat{\mu}} - \hat{\mu}(R+v_\alpha(i'))] = \lim_{\alpha\to 0^+}[c_{\tilde{\mu}} - \tilde{\mu}(R+v_\alpha(i'))],$$

so that $v_\alpha(i')$ strictly decreasing in α yields the desired result.

Q.E.D.

In conclusion, we note a rather curious phenomenon. From (18) it is apparent that $v_\alpha(i) \geq hi[\Lambda/(\alpha+\Lambda)]^i/\Lambda$, so that

$$h(i) = \lim_{\alpha \to 0^+} [V_\alpha(i) - V_\alpha(0)] = \lim_{\alpha \to 0^+} \sum_{j=1}^{i} v_\alpha(j) \geq \frac{h}{\Lambda} \sum_{j=1}^{i} j = \frac{h}{\Lambda} \frac{i(i+1)}{2} .$$

Thus, although the reward function is linearly bounded, the relative values $h(\cdot)$ are quadratic.

VI. AN M/M/c WITH VARIABLE ARRIVAL RATE

Closely related to Crabill's M/M/1 model with variable service rate is Low's [13,14] M/M/c model with variable arrival rate. In contrast to Crabill's model, control of the system is effected by increasing or decreasing the price charged for the facility's service thereby encouraging or discouraging the arrival of customers. This scenario is, of course, equivalent to the decision maker choosing an arrival rate λ which in turn determines the price p_λ that a customer is charged for admission to the queue.

We assume that the arrival rate lies in some closed subset A of $[0,\bar{\lambda}]$, $\bar{\lambda} < \infty$, each of the $c < \infty$ exponential servers has rate μ, and the queue capacity Q is allowed to be either finite or infinite.

The state of the system is merely the number of customers in the system, and the cost structure consists of two parts: a holding cost h_i per unit time that the system is in state i and a reward or entrance fee p_λ received whenever a customer enters the system at a point in time when the arrival rate is λ. As is reasonable from economic considerations, p_λ is taken to be nonincreasing whereas h_i is nondecreasing. In addition, p_λ is assumed to be right-continuous, with $p_{\bar{\lambda}} \geq 0$ and $p_0 < \infty$. Unlike Low, we must assume that h_i is a convex function of i. Setting $\Lambda = \bar{\lambda}+\mu c$, we obtain the recursive equations $(V_0 \equiv 0)$

(21) $$V_{n+1,\alpha}(i) = \frac{1}{\alpha+\Lambda} \min_{\lambda \varepsilon A} \{-\lambda p_\lambda + V_{n,\alpha}(i,\lambda)\}, \quad i = 0,1,\ldots,c+Q \ ,$$

where

(22) $$V_{n,\alpha}(i,\lambda) = h_i + \lambda V_{n,\alpha}(i+1) + \mu(i \wedge c)V_{n,\alpha}(i-1) + (\Lambda-\lambda-\mu(i \wedge c))V_{n,\alpha}(i) \ .$$

As before, we begin by defining $\lambda_{n,\alpha}(i)$ to be an optimal arrival rate when the system is in state i, n periods remain, and the discount factor is $\alpha \geq 0$. Also define

$$v_{n,\alpha}(i) = V_{n,\alpha}(i) - V_{n,\alpha}(i-1) .$$

Clearly, $v_{n,\alpha}(i) \geq 0$.

Low's principal thrust was the development of an efficient algorithm for the average cost case from which optimality of a monotone stationary policy was established. We now extend his monotonicity result to the finite and infinite horizon discounted problem.

<u>THEOREM 18</u>. For each n and $\alpha \geq 0$, $V_{n,\alpha}(\cdot)$ is convex, so that $\lambda_{n,\alpha}(i)$ is a nonincreasing function of i.

<u>Proof</u>: Convexity of $V_{1,\alpha}$ follows from that of h_i. Assume $V_{n,\alpha}$ is convex. Then using convexity of V_n and the method of Theorem 11, we obtain, after some simplification,

$$\begin{aligned}(\alpha+\Lambda)[v_{n+1}(i+1)-v_{n+1}/(i)]/\mu \\ \geq c[v_n(i+1)-v_n(i)] - \{((i+1)\wedge c)v_n(i+1) - 2(i\wedge c)v_n(i) \\ +((i-1)\wedge c)v_n(i-1)\} .\end{aligned}$$

The nonnegativity of this last expression is shown by considering the three cases $i+1 \leq c$, $i = c$, and $i-1 = c$ separately. In each case, the desired result follows from convexity of V_n. This completes the induction argument.

Rewriting (21) as

$$(23)\quad V_{n+1,\alpha}(i) = \frac{1}{\alpha+\Lambda}\{h_i+\Lambda V_{n,\alpha}(i)-\mu(i\wedge c)v_{n,\alpha}(i) + \min_{\lambda\varepsilon A}\{\lambda[v_{n,\alpha}(i+1)-p_\lambda]\}\} ,$$

we see that right-continuity of p_λ together with A closed, $v_{n,\alpha}(i+1) \geq 0$, and p_λ nonincreasing guarantees the existence of a minimizing λ, whereas $\lambda_{n+1,\alpha}(i)$ nonincreasing in i is obtained from $v_{n,\alpha}(i)$ nondecreasing in i.

Q.E.D.

<u>REMARK</u>. Equation 23 reveals that if p_λ is continuous, then $\lambda_{n+1,\alpha}(i)$ is nonincreasing in i whether or not p_λ is nonincreasing.

Letting $\overline{\lambda}$ play the role of $\overline{\mu}$ in the proofs of Theorems 12 and 14 and letting $\underline{\lambda} \equiv \inf\{\lambda:\lambda\varepsilon A\}$ play the role of $\overline{\mu}$ in the proof of Lemma 4, the proofs of Theorems 12, 13, 14, 15 and Lemma 4 suffice, with but minimal changes, to establish the obvious analogs of Theorems 12, 13, 14, 15 and Lemma 4 for the variable arrival rate model. Assuming that $\underline{\lambda} < c\mu$ rather than $\overline{\lambda} < c\mu$ if $Q = \infty$, the analog of Theorem 16 holds. Finally, the partial ordering on the set of 0-optimal stationary policies (cf. Theorems 10 and 17) relates that faster arrival rates are preferred.

REFERENCES

1. Bell, C., "Characterization and Computation of Optimal Policies for Operating an M/G/1 Queuing System with Removable Server," Opns. Res., 19 (1971), 208-218.

2. Bell, C., "Optimal Average-Cost Operating Policy for an M/G/1 Queuing System with Removable Server and Several Priority Classes," Technical Report No. 29, Department of Operations Research, Stanford University, 1971.

3. Blackburn, J., "Optimal Control of a Single-Server Queue with Balking and Reneging," Management Science, 19 (1972), 297-313.

3a. Blackwell, D., "Discrete Dynamic Programming," Ann. Math. Stat., 33 (1962), 719-726.

4. Cinclar, E., Introduction to Stochastic Processes, Prentice Hall, forthcoming 1974.

5. Crabill, T., "Optimal Control of a Queue with Variable Service Rates," Technical Report No. 75, Department of Operations Research, Cornell University, Ithaca, New York, 1969.

6. Crabill, T., "Optimal Control of a Service Facility with Variable Exponential Service Times and Constant Arrival Rate," Management Science, 18 (1972), 560-566.

7. Cramer, M., "Optimal Customer Selection in Exponential Queues," ORC 71-24, Operations Research Center, University of California, Berkeley, 1971.

8. Denardo, E. V., "Markov Renewal Programs with Small Interest Rates," Ann. Math. Stat., 42 (1971), 477-496.

9. Farrell, W., "Optimal Control of the M/M/c Queue with Removable Servers," Working Paper, Western Management Science Institute, UCLA, forthcoming.

10. Heyman, D., "Optimal Operating Policies for M/G/1 Queuing Systems," Opns. Res., 16 (1968), 362-382.

11. Lippman, S., "Semi-Markov Decision Processes with Unbounded Rewards," Management Science, 19 (1973), 717-731.

12. Lippman, S. and S. Ross, "The Streetwalker's Dilemma: A Job-Shop Model," SIAM J. Appl. Math., 20 (1971), 336-342.

13. Low, D., "Optimal Dyanmic Operating Policies for an M/M/s Queue with Variable Arrival Rate," Working Paper No. 178, Western Management Science Institute, UCLA, August 1971.

14. Low, D. W., "Optimal Dynamic Pricing Policies for an M/M/s Queue," Opns. Res., to appear.

15. Magazine, M., "Optimal Control of Multi-Channel Service Systems," Nav. Res. Log. Quart., 18 (1971), 177-183.

16. McGill, J., "Optimal Control of a Queuing System with Variable Number of Exponential Servers," Technical Report No. 123, Department of Operations Research and Department of Statistics, Stanford University, 1969.

17. Miller, B., "Finite State Continuous-Time Markov Decision Processes with Applications to a Class of Optimization Problems in Queueing Theory," Technical Report No. 15, Program in Operations Research, Stanford University, 1967.

18. Miller, B., "A Queuing Reward System with Several Customer Classes," Management Science, 16 (1969), 234-245.

18a. Miller, B. and A. F. Veinott, Jr., "Discrete Dynamic Programming with a Small Interest Rate," Ann. Math. Stat., 40 (1969), 366-370.

19. Prabhu, N. and S. Stidham, Jr., "Optimal Control of Queueing Systems," Technical Report No. 186, Department of Operations Research, Cornell University, Ithaca, New York, 1973.

20. Sabeti, H., "Optimal Decision in Queueing," ORC 70-12, Operations Research Center, University of California, Berkeley, 1970.

21. Sobel, M., "Optimal Average-Cost Policy for a Queue with Start-Up and Shut-Down Costs," Opns. Res., 17 (1969), 145-162.

22. Veinott, A. F., Jr., "On Finding Optimal Policies in Discrete Dynamic Programming with No Discounting," Ann. Math. Stat., 37 (1966), 1284-1294.

23. Yadin, M. and S. Zacks, "The Optimal Control of a Queueing Process," Operations Research, Statistics, and Economics Mimeograph Series No. 54, Technion, Haifa, Israel, 1969.

24. Zacks, S. and M. Yadin, "Analytic Characterization of the Optimal Control of a Queueing System," J. Appl. Prob., 7 (1970), 617-633.